

CLAIMS

What is claimed is:

- 5 1. A method for asynchronously tracking the replication of data writes, from an application that is subject to system and network failure, to a standby data processing unit located at a standby site, the method comprising the steps of:

 - (a) enabling access to sequence numbers created for use in replication of the data writes;
 - 10 (b) sending the data writes from an application write to a remote mirroring module;
 - (c) assigning sequence numbers to each data write accepted;
 - (d) tracking a most recent local write sequence number for a local data write and a most recent replication sequence number for replicated data writes;
 - 15 (e) determining when data writes at a local site have been successfully replicated at the standby site by correlating the most recent local write sequence number and most recent replication sequence number; and
 - (f) initiating replication of the data writes for which replication has not taken place.
- 20 2. A method in accordance with claim 1, further comprising the step of comparing the most recent local write sequence number and most recent replication sequence number to identify a group of data writes to be re-sent through the remote mirroring module.
- 25 3. A method in accordance with claim 1, further comprising the step of beginning the replication recovery with the data write immediately following the most recent replication sequence number that has been recorded.
- 30 4. A method in accordance claim 1, further comprising the step of resubmitting replication requests through the remote mirroring system for each data write that has a sequence number higher than the most recent replication sequence number and less than or equal to the most recent local write sequence number.

5. A method in accordance with claim 1, further comprising the step of determining when to report that a specific application transaction has been reliably replicated based on the most recent replication sequence number tracked for data blocks that have been replicated.
6. A method in accordance with claim 1, further comprising the step of performing in-order mirror resynchronization by replicating data blocks with sequence numbers that are greater than the most recent replication sequence number.
7. A method in accordance claim 1, further comprising the step of using the most recent replication sequence number to determine the last data write replicated.
8. A method for handling overflow of disk spool writes in a remote mirroring unit subsequent to a network or system failure, comprising the steps of:
 - (a) sending data writes that need to be replicated from the application through the remote mirroring unit which includes a disk spool;
 - (b) recording data writes that are not spooled in the disk spool due to a spool overflow, wherein the unspooled data writes are recorded in a spool overflow list; and
 - (c) using the spool overflow list to enable subsequent resynchronization of a standby storage system with the primary storage system for the application.
9. A method in accordance with claim 8, further comprising the step of storing the spool overflow list of data blocks that were not sent to the disk spool in a chunk map area, when the disk spool cannot accept further input.
10. A method in accordance with claim 8, further comprising the step of storing the spool overflow list of the data blocks that were not sent to disk spool in a bit map to guide further resynchronization of the standby storage system with the application's primary storage system.
11. A method in accordance with claim 9, further comprising the step of resynchronizing the standby storage system with the application's primary storage system using the bit map to determine which data blocks need to be replicated.

12. A method in accordance with claim 9, further comprising the step of performing synchronization of the standby storage system with the application's primary storage system using the bit map.

5

13. A method in accordance with claim 9, further comprising the step of using a chunk map area as the bit map, wherein each chunk map element maps to a storage chunk that can have data blocks written into the chunk.

10

14. A method in accordance with claim 8 further comprising the step of reversing the roles of the standby storage system and the primary storage system, wherein a new standby storage system is synchronized with a new primary storage system using information in the spool to minimize the amount of information that is shipped between primary storage system and standby storage system.

15

15. A device for handling disk spool overflow in a remote mirroring system, comprising:

(a) a data transaction application that includes a plurality of data blocks that are intended to be replicated in a mirrored storage subsystem;

20

(b) a primary remote mirroring module having a local disk spool, configured to send data blocks received from the data transaction application to a standby data processing system for replication;

(c) a standby remote mirroring module, operating in the standby data processing system, including a standby disk spool to receive the data blocks sent by the primary remote mirroring module; and

25

(d) a spool overflow in the primary remote mirroring module, including a memory cache, for recording writes of data blocks when the local disk spool has overflowed.

30

16. A device as in claim 15, further comprising a bit map that is associated with the spool overflow in order to track unspooled data block writes.

17. A device in accordance with claim 15, further comprising cache memory associated with the primary remote mirroring system, wherein the cache memory is used to store data blocks prior to transfer of the data blocks to the disk spool.
- 5 18. A device in accordance with claim 15, wherein the primary remote mirroring module and the standby remote mirroring module are both software programs.
- 10 19. A method for decreasing the number of writes to a disk spool in a remote mirroring unit in order to reduce the amount of disk spool input and output (I/O) taking place, wherein the remote mirroring unit supports an application having a plurality of data blocks stored in a primary storage system, comprising the steps of:
- (a) sending data blocks that need to be replicated from the application to the remote mirroring unit which includes a disk spool;
 - 15 (b) delaying data blocks writes to the disk spool;
 - (c) writing data blocks to the disk spool when a conflicting write request comes to the same data block; and
 - (d) replicating the data blocks to resynchronize a standby storage system with the primary storage system for the application.
- 20 20. A method in accordance with claim 19, further comprising the step of minimizing spool writes by holding data block writes in a cache located in the remote mirroring unit.
- 25 21. A method in accordance with claim 19, further comprising the step of storing writes of the data blocks in a cache until a conflicting transaction write triggers transfer of the cached data blocks to the disk spool.
- 30 22. A method in accordance with claim 19, further comprising the step of storing writes of the data blocks in a cache until memory constraints trigger the transfer of the cached data blocks to the disk spool.

23. An article of manufacture, comprising:

a computer usable medium having computer readable program code means embodied therein for asynchronously tracking the replication of data writes, from an application that is subject to system and network failure, to a standby data processing unit located at a standby site, the computer readable program code means in the article of manufacture comprising:

- (a) computer readable program code means for enabling access to sequence numbers created for use in replication of the data writes;
- (b) computer readable program code means for sending the data writes from an application write to a remote mirroring module;
- (c) computer readable program code means for assigning sequence numbers to each data write accepted;
- (d) computer readable program code means for tracking a most recent local write sequence number for a local data write and a most recent replication sequence number for replicated data writes;
- (e) computer readable program code means for determining when data writes at a local site have been successfully replicated at the standby site by correlating the most recent local write sequence number and most recent replication sequence number; and
- (f) computer readable program code means for initiating replication of the data writes for which replication has not taken place.

24. An article of manufacture, comprising:

a computer usable medium having computer readable program code means embodied therein for handling overflow of disk spool writes in a remote mirroring unit subsequent to a network or system failure, the computer readable program code means in the article of manufacture comprising:

computer readable program code means for sending data writes that need to be replicated from the application through the remote mirroring unit which includes a disk spool;

computer readable program code means for recording data writes that are not spooled in the disk spool due to a spool overflow, wherein the unspooled data writes are recorded in a spool overflow list; and

computer readable program code means for using the spool overflow list to enable subsequent resynchronization of a standby storage system with the primary storage system for the application.

- 5 25. A device for asynchronously tracking the replication of data writes, from an application that is subject to system and network failure, to a standby data processing unit located at a standby site, comprising:
- (a) means for enabling access to sequence numbers created for use in replication of the data writes;
 - 10 (b) means for sending the data writes from an application write to a remote mirroring module;
 - (c) means for assigning sequence numbers to each data write accepted;
 - (d) means for tracking a most recent local write sequence number for a local data write and a most recent replication sequence number for replicated data writes;
 - 15 (e) means for determining when data writes at a local site have been successfully replicated at the standby site by correlating the most recent local write sequence number and most recent replication sequence number; and
 - (f) means for initiating replication of the data writes for which replication has not taken place.
- 20 26. A device as in claim 25, further comprising a means for comparing the most recent local write sequence number and most recent replication sequence number to identify a group of data writes to be re-sent through the remote mirroring module.
- 25 27. A device for handling overflow of disk spool writes in a remote mirroring unit subsequent to a network or system failure, comprising:
- (a) means for sending data writes that need to be replicated from the application through the remote mirroring unit which includes a disk spool;
 - 30 (b) means for recording data writes that are not spooled in the disk spool due to a spool overflow, wherein the unspooled data writes are recorded in a spool overflow list; and

(c) means for using the spool overflow list to enable subsequent resynchronization of a standby storage system with the primary storage system for the application.

5 28. A device as in claim 27, further comprising means for storing the spool overflow list of data blocks that were not sent to the disk spool in a chunk map area, when the disk spool cannot accept further input.

10 29. A device as in claim 27, further comprising means for storing the spool overflow list of the data blocks that were not sent to disk spool in a bit map to guide further resynchronization of the standby storage system with the application's primary storage system.

10099760-034502